

A Novel Method for Characterizing Spacesuit Mobility Through Metabolic Cost

Shane M McFarland¹ and Jason R. Norcross²
Wyle/NASA-Johnson Space Center, Houston, TX, 77058

Historically, spacesuit mobility has been characterized by directly measuring both range of motion and joint torque of individual anatomic joints. The work detailed herein aims to improve on this method, which is often prone to uncertainty, lack of repeatability, and a general lack of applicability to real-world functional tasks. Specifically, the goal of this work is to characterize suited mobility performance by directly measuring the metabolic performance of the occupant. Pilot testing was conducted in 2013, employing three subjects performing a range of functional tasks in two different suits prototypes, the Mark III and Z-1. cursory analysis of the results shows the approach has merit, with consistent performance trends toward one suit over the other. Forward work includes the need to look at more subjects, a refined task set, and another suit in a different mass/mobility regime to validate the approach.

Nomenclature

<i>BTU</i>	=	British thermal unit
<i>CO₂</i>	=	carbon dioxide
<i>EVA</i>	=	extravehicular activity
<i>Kcal</i>	=	kilocalories
<i>kg</i>	=	kilogram
<i>L</i>	=	liters
<i>L/min</i>	=	liters per minute
<i>O₂</i>	=	oxygen
<i>VCO₂</i>	=	rate of CO ₂ production
<i>VO₂</i>	=	rate of O ₂ consumption

I. Introduction

Spacesuit mobility has historically been defined and characterized by a combining range of motion analysis with joint torque measurement of individual anatomical joints performing isolated motions meant to drive that joint only in a given orthogonal plane.^{1,2,3} Although this has been the standard approach for several decades, there are numerous shortcomings. First, the lack of a standardized method for collecting both range of motion and joint torque translates to many different test setups, procedures, and methods of data analysis.^{1,3} Second, all of these previously used methods for data collection demonstrate a lack of some degree of repeatability, even within the same test setup and the same conductor.³ For example, the standard fish-scale method has been used for numerous range-of-motion and joint torque tests. The results show high variability even within one test point of multiple joint articulations – much less the variability seen using a different fish scale or different test conductor.⁴ In addition, attempts at higher fidelity data collection techniques, such as motion capture, require high overhead and cost with minimal improvement.^{1,5} Lastly, and perhaps most importantly, isolated motions in standard anatomical planes are not representative of real-world tasks that a crewmember would be performing during an extravehicular activity (EVA), be it microgravity or surface exploration based.⁶

To address these shortcomings, NASA is exploring options to ascertain the feasibility of alternative approaches to defining mobility – methods that are more repeatable, lower overhead, and more tied to functional EVA tasks. One such option looks at the metabolic energy-cost of a spacesuit. In other words, can one objectively compare the mobility of a spacesuit by evaluating the metabolic cost of that suit to the wearer? This attempts to address the issue of spacesuit mobility not at the individual joint level, but of the overall suit system while performing representative EVA tasks.

¹ Senior Spacesuit Project Engineer, Spacesuit and Crew Survival Systems Branch

² Lead Scientist, EVA Physiology Laboratory, Wyle Science, Technology and Engineering Group.

II. Methods and Approach

As this study was truly a feasibility assessment of the approach, and not an attempt at specifically characterizing mobility of one suit versus another, a balance was attempted between collecting as much data in as many different ways as possible, collecting enough data to be statistically relevant, and total test overhead. As a result, the following test parameters were selected, which represent both standard functional tasks performed in routine suited fit-checks, and tasks meant to exercise all the significant joints.

- Three subjects – all experienced suited subjects
- Two suits – Mark III and Z-1 planetary prototype
- Eight functional tasks performed for 2 minutes each
 - 1) Walk: Walk a 14' straight path; turn around and repeat
 - 2) Sit: Sit onto a 14" tall box; stand and repeat
 - 3) Climb: Step up onto two 7" tall carpeted steps in sequence; step backwards down to the floor, and repeat
 - 4) Prone and recover: From a standing position, move to a prone position, stand, and repeat
 - 5) Shovel: Move cork mulch from one 45" x 48" x 9" box to an adjacent 50 L plastic container using a standard spade shovel weighted with a 5-lb sand bag at the junction of the spade and handle
 - 6) Hammer: Using a rubber mallet, strike 7" square corrugated cardboard pads at waist height, one in the vertical plane (30" height) and the other in the horizontal plane (39" height)
 - 7) Object relocation: Move a 10-lb medicine ball from a waist-height shelf 42" tall to the floor 36" away from the base of the shelf, back to the shelf, and repeat
 - 8) Side step: Step sideways onto a 7" carpeted step, to floor on other side, and repeat
- A minimum of 2 minutes of seated rest was enforced between each task
- If the subject was not below a metabolic rate of 1000 British thermal units (BTU)/hr after 2 minutes, additional resting time was provided

For each subject, the battery of tasks was performed four ways:

- Mark III Suit
- Z-1 suit
- Unsuit, at a pace that felt "natural" to the subject
- Unsuit, at a pace and cadence that matched their previous Mark III suited run

Prime collected data included: carbon dioxide (CO₂) production, time to completion, task repetitions, and subjective rate of perceived exertion. Secondary data of interest that were collected included: heart rate, skin temperature, and breathing rate. Note that although these secondary data were collected in the event it was potentially helpful, it was not intended for specific analysis.

Each subject's first run was always unsuited (natural). The second run was always suited – either Mark III or Z-1, depending on subject. The third run was also suited, but the subject was told to continue past the 2 minutes to match the same number of repetitions as the first suit run, if necessary. The fourth and final run was always unsuited, with a matched pace to that of the Mark III run. Doing the test in this manner allowed the test conductors to later determine which metrics were the best for comparison, while having all the data necessary to compare one run against the other. Each subject did all test runs in one day, with a large break between the two suited runs to offset fatigue to the extent possible.

It should be noted that all aspects of the functional tasks (distance between items, size, weight, etc.) were collected, not only to ensure consistency between subjects and test runs, but also to potentially use as a later standard. It should also be mentioned that two minutes was selected as the candidate time for each task based on previous testing performed at NASA collecting metabolic data and the desire to keep this feasibility assessment within reasonable testing limits. That said, many recommendations for forward work, including time for each task are documented in Section VII of this paper.

III. Test Article Description

For data collection purposes, suited energy expenditure was based on measured inlet suit ventilation rate and expired CO_2 concentration in the exhaust umbilical (via a CD-3A Infrared Carbon Dioxide Analyzer, AEI Technologies, Pittsburgh, PA). This information is used to calculate a rate of CO_2 production (VCO_2). A constant respiratory exchange ratio of 0.85 is assumed to estimate the rate of oxygen consumption (VO_2). This technique and hardware have been described in previously^{7,8,9}. We assume that the ventilation rate of 6 ACFM and directional flow of air over the face ensures proper gas mixing throughout the suit and exhaust umbilical and any leaks occurring from the suit have the same gas composition as measured in the exhaust umbilical.

Although this data collection method limits the ability to collect breath by breath CO_2 generation, it was not felt to be an issue, given the fact that the tasks are being performed for 2 minutes. This test was not attempting to gather instantaneous metabolic load, but rather the average metabolic load over the entire 2-minute task duration. The CO_2 analyzer was assembled per NASA drawing A16-M00020, Rev I.

A COSMED K4b2 mobile metabolic system was worn by the subjects for unsuited data collection. This system consists of a small chest or back-worn box and an oral-nasal mask. In addition, heart and respiration rates and skin temperature were collected by a Zephyr BioHarness for suited as well as unsuited runs. Both of these items are commercial off-the-shelf products.

Unsuited energy expenditure was measured by the COSMED K4b2 with breath-by-breath measurement of VCO_2 and VO_2 . Conversions to kcal/hr and BTU/hr remain the same.

The CO_2 analyzer system and COSMED K4b2 are shown in Figures 1 and 2, respectively.



Figure 1. COSMED K4b2.

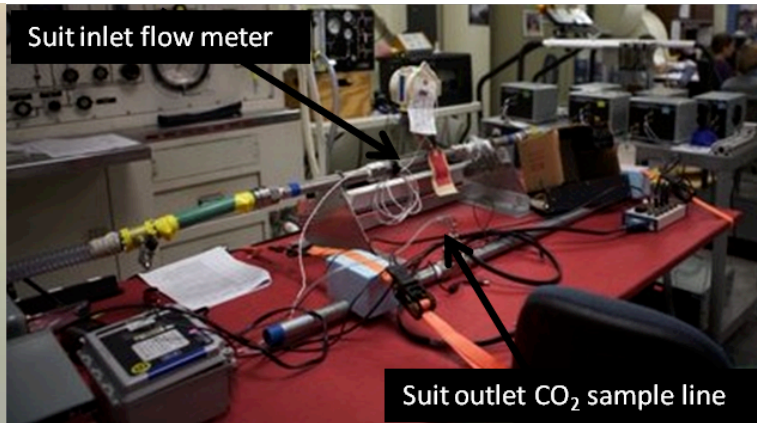


Figure 2. CO_2 analyzer system.

IV. Test Anomalies

Several anomalies and adjustments occurred through the course of testing, as should always be expected. Although the test was mostly completed as defined above, this section serves to document those anomalies for preservation and consideration by the reader.

- Subject B was unable to perform the prone and recover task, and communicated that to the test conductor at the beginning of the test day. Therefore, the decision was made to have Subject B perform a kneel and recover task, with the aid of a walking stick, for all runs – suited and unsuited – to allow for the best comparison. Note that this means that future results in this paper present Subject B as performing a prone/recover task when, in reality, it was a kneel and recover task. However, it was determined that this was acceptable, given the “feasibility assessment” goal of this effort, and that direct comparison between subjects was not a component of analysis.
- Subject C experienced a seal leak on the Mark III suit about halfway through the run, 90 seconds into the prone and recover task. This was Subject C’s first suited run. The decision was made to not make up this dropped data, as the testing was a feasibility assessment only and there would have been a delay of at least 1 day. The decision was made to collect all Z-1 data, with the first half matching the Mark III (including the shortened 90-second prone and recover task), and the second half being the standard 2-minute durations as if it were the first suited run.

Other than these two items, the test was completed as specified.

V. Analysis

The collected data were analyzed using two different metrics: First, O_2 consumption was calculated as a total cost defined as liters (L) of O_2 per task or as a consumption rate in L/min. Because these were not steady state tasks of consistent workload, the standard VO_2 data in L/min is of little value. But although the total O_2 cost of the task allows for high level comparison, it does not account for possible differences in total repetitions performed and suit mass. Secondly, these data were normalized to two factors – mass of the subject and suit system, and number of repetitions. As such, the units for this second metric were ml O_2 consumed per kg combined mass per rep (ml/kg/rep). Although these spacesuits were of similar mass, the addition of suit mass to the normalization scheme may not be relevant, but in future evaluations that included suits with different masses will need to consider how that affects performance.

Given only two complete data sets (as described in Section IV) and that the fact that this was a pilot study with an unevaluated protocol, the caveat must be made that the results as presented here are meant to ascertain feasibility – not compare the two suits against each other. Future testing will be required for a statistically relevant comparison based on lessons learned from this testing including repeated measures of the same subject in the same spacesuit to determine within subject variability and to be able to determine if the differences between performance across spacesuits is greater than the known variability within the same spacesuit architecture.

VI. Results

One initial finding is that for many tasks, the subjects were just reaching or never reached a steady state metabolic rate. Individual subjects vary in speed of metabolic response. To keep the overall test shorter, we selected a duration of 2 minutes because we felt this would be the minimum duration to show possible metabolic differences between tasks. While we were able to see differences with 2-minute durations, which will be discussed following, we believe that extending the duration of these tasks would more clearly highlight these differences.

There are multiple ways of displaying the same data for comparison. After some analysis, we determined stacking bar charts to be the most logical and informative. These charts can be organized by stacking tasks on top of each other, or by stacking the difference (absolute or percent) between the two suits for all subjects. In addition, a weighting factor can be applied to value some tasks as more frequent during an EVA than others, thus affecting the overall value of specific individual tasks.

Figure 3 (top) shows the total O_2 consumed per each task and overall O_2 consumed for the whole task battery. Subject C did not complete the last 4 tasks during the MKIII test point. No one task dominates the total O_2 consumption, but this figure does not account for possible variation in the total number of repetitions completed per task.

Figure 4 (bottom) normalizes this total expenditure to the combined mass of the suit and subject as well as the number of repetitions. Normalizing to each repetition dramatically increased the overall weight of certain tasks such as prone/recover and object relocation and reduced the relative contribution of high-repetition tasks, such as hammering and shoveling.

While interesting, neither of these viewpoints effectively describes the overall data. Figure 4 (top) would be a very effective way to compare the overall differences between two suits if total overall workload was controlled. In this feasibility study, we primarily controlled for time. Normalizing to repetitions allows for effective comparison of suited performance within the same task and subject.

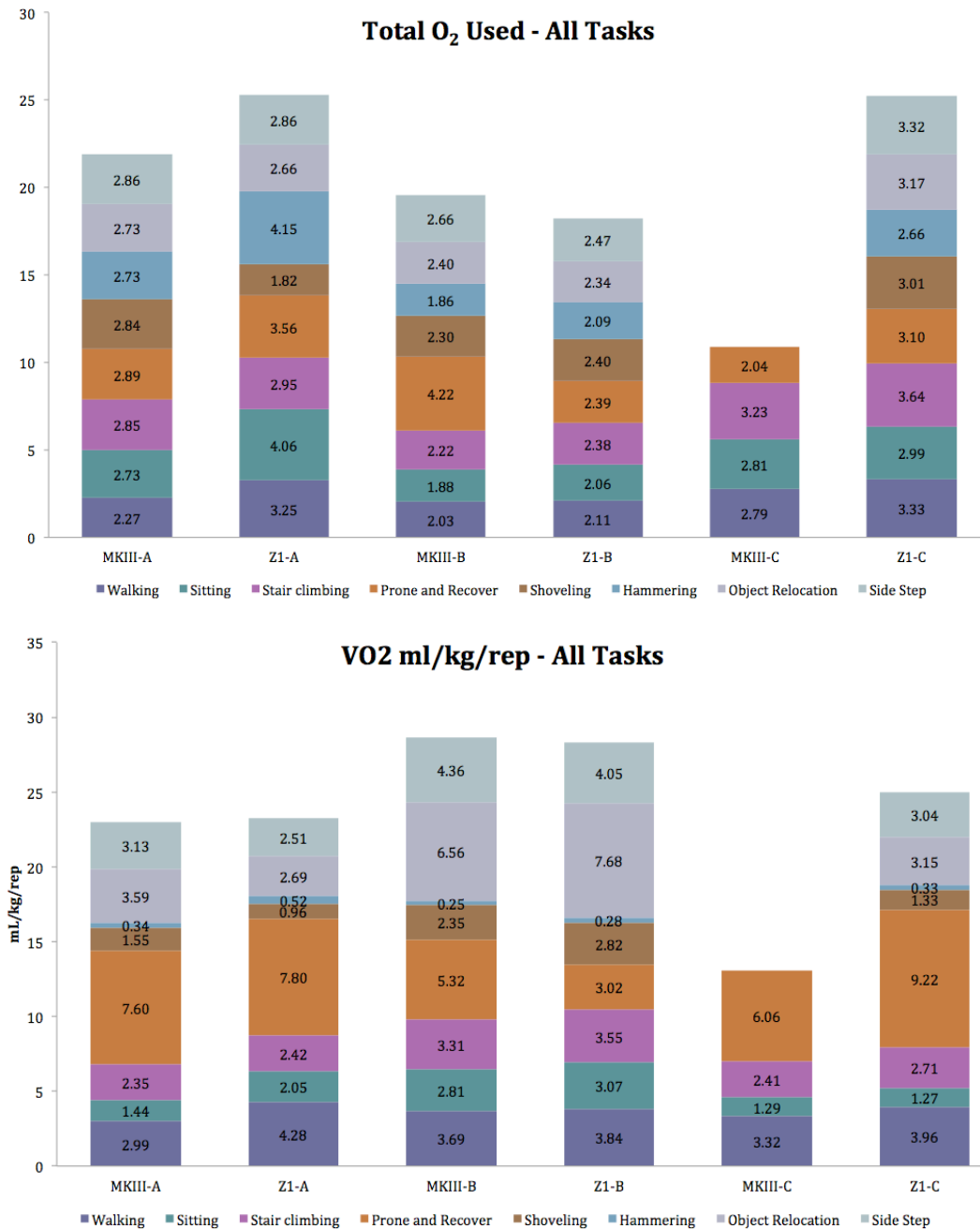


Figure 3. Stacking bar charts, showing total L O₂ consumed per task (top) and mass/rep normalized VO₂ along the vertical axis as each subject (A, B, C) completes the various tasks in each suit (MKIII, Z1)

The primary objective was to determine if metabolic differences between task performance existed while wearing different EVA suits. Figure 4 (top) shows the difference in total O₂ consumption between the Mark III and Z-1 for each subject. In most cases, the trend shows that subjects required less energy to complete tasks in the Mark III, but without equal workloads, this chart could be misleading. Figure 4 (bottom) shows the normalized data per each repetition. With subjects completing a similar number of repetitions, the differences between the absolute and normalized data were not immediately noticeable.

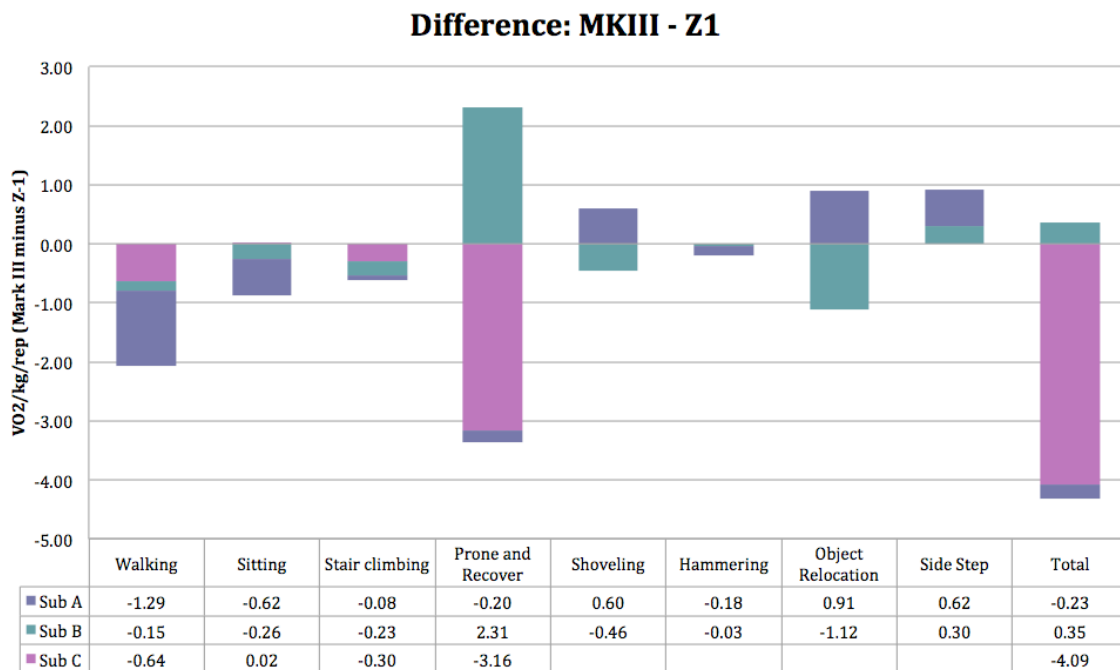
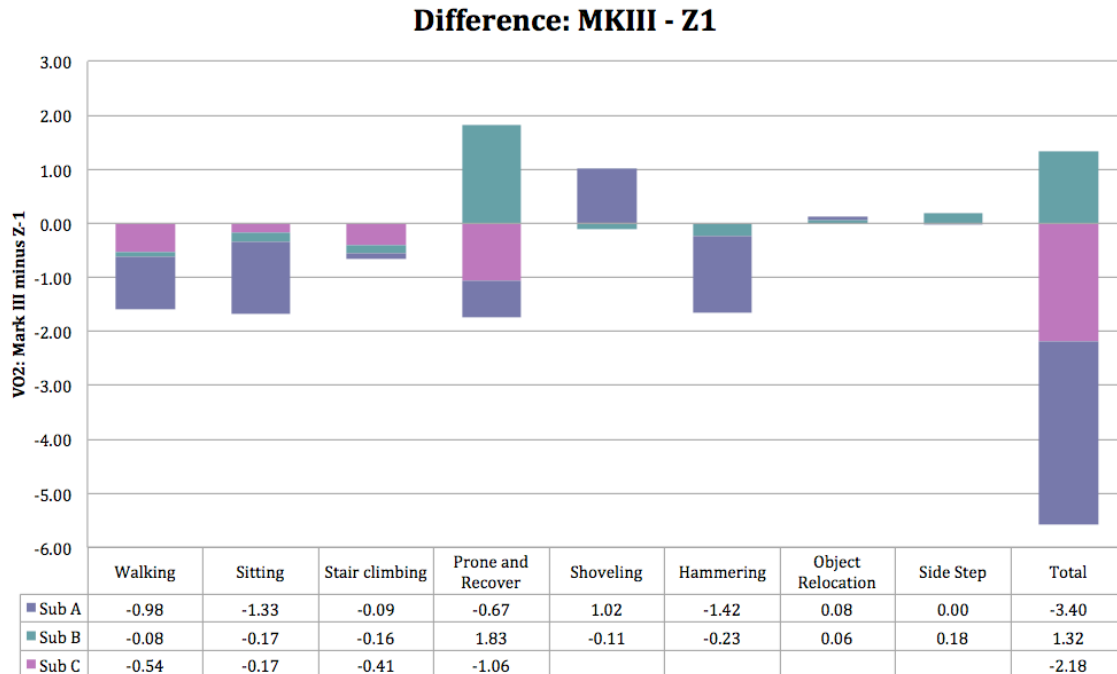


Figure 4. Metabolic differences between the two suits for each task with all subjects stacked together. Negative values designate a metabolic advantage for the Mark III; positive numbers designate a Z-1 advantage. When all data is positive or negative, such as in the walking task, it designates a unanimous performance advantage trend toward one suit or the other.

Figure 5 takes the same data from Figure 4, but shows the percent difference between the two suits for each task, and subjects stacked together. Similar positive/negative designations and unanimous trends can be observed as in Figure 4 with the absolute differences. The main advantage with percent difference is that we could possibly prospectively define a significant difference for future suited evaluation of mobility, especially with further testing combined with sensitivity analysis based on consumables usage.

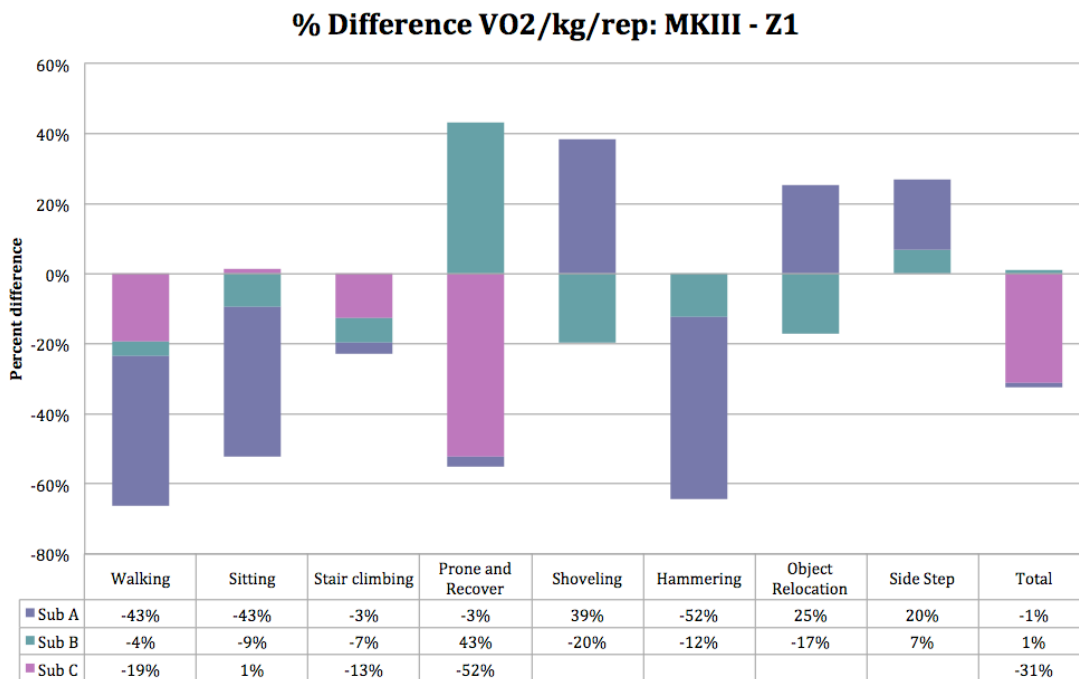
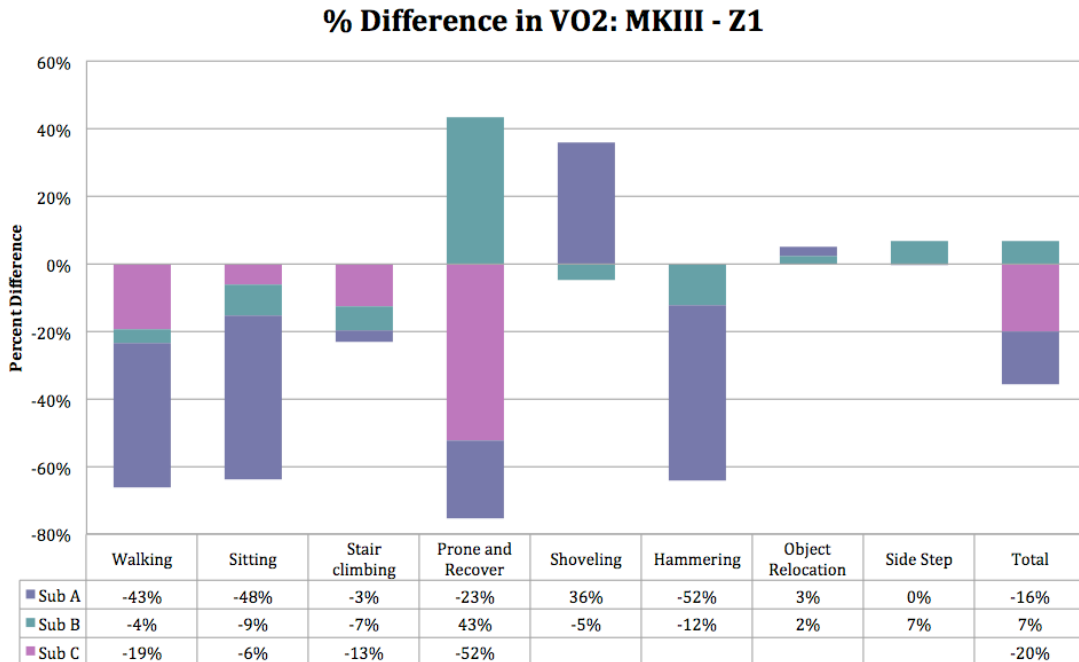


Figure 5. Percent difference for metabolic variables between the two suits for each task with all subjects stacked together. Negative values designate a metabolic advantage for the Mark III; positive numbers designate a Z-1 advantage.

A preliminary attempt was made at weighting the tasks based on assumptions defined in ¹⁰NASA TM-2010-216138, "Life Sciences Implications of Lunar Surface Operations." This weighting prioritized the tasks based on expected surface EVA operations, where walking may be more prevalent than climbing, for instance. Most tasks did not change weight appreciably, with the exception of walking, which increased from 12.5% to 30%, and sitting, which decreased from 12.5% to 1%. The detailed results of this type of analysis would be premature with the limited

data set, but is mentioned in this paper to demonstrate how varying EVA concepts of operation could be evaluated to ascertain which suit might offer better human performance.

Upon detailed study of these results, it becomes evident that some significant trends toward one suit in many tasks exists, with perhaps a smaller trend toward the other suit on another task. Results were tabulated for the two metrics (total O₂ consumed and normalized VO₂ [mL/kg/rep]) for all tasks and subjects in an attempt to display all of these results in one take-away “bingo chart.” The results are shown in Figure 6, with dark colors representing a difference of 10 percentage points or more, and lighter colors representing a difference of less than 10 percentage points.

		Total O ₂		Total O ₂ <10%		VO ₂ (mL/kg/rep)		VO ₂ (mL/kg/rep) <10%	
		Dark		Light		Dark		Light	
		Walking	Sitting	Stair Climbing	Prone/Kneel and Recover	Shoveling	Hammering	Object Relocation	Side Step
Mark III	A	Dark	Dark	Light	Dark		Dark		
	B	Light	Light	Light		Light	Dark	Light	
	C	Dark	Light	Dark	Dark				
Z-1	A					Dark		Light	Light
	B				Dark			Light	Light
	C		Light						

Figure 6. Tabulated results for all metrics, subjects and suits.

As can be clearly shown, there are some consistent, unanimous trends toward the Mark III suit being “better” (i.e., less metabolic cost) on tasks such as walking and stair climbing, and another unanimous trend toward the Z-1 suit in the side-step task.

The most consistent performance trends were as follows:

- Mark III bests Z-1 in walking task (20+% unanimous)
- Mark III bests Z-1 in stair climbing task (8% unanimous)
- Mark III bests Z-1 in sitting task (19% not unanimous)
- Mark III bests Z-1 in hammering task (32% unanimous, n=2)
- Z-1 bests Mark III in side step task (9% unanimous, n=2)

The most “mixed” trends include:

- Prone/kneel and recover
- Object relocation
- Shoveling

Although interesting pilot data, there are major flaws that need to be addressed before any actual comparisons can be made. We currently have no repeated measures data to characterize the variability seen with the same subject in the same suit repeating these tasks many times. It could be that the metabolic differences seen between performance in both suits is within the normal range of variability for one suit. Until this factor is understood, we cannot make any solid claims about performance within a given suit.

Note, again, this testing was trying to answer the primary question, “Can we objectively compare the mobility of a spacesuit by evaluating the metabolic cost of that suit to the wearer while performing a battery of functional EVA tasks?” Based on the results as outlined here, it appears that the answer is clearly “Yes”; however, there were several aspects of the method and protocol noted during testing that, if improved, should yield an even clearer picture and allow us to vastly improve the fidelity of the approach.

VII. Recommendations

The feasibility of this approach for characterizing and comparing the mobility of space suit configurations using the energy-based method was demonstrated.¹¹ Whether this approach can be improved and fully developed into a standardized, repeatable method is yet to be seen. Based on this pilot test, several aspects of the protocol invite improvements. It is felt that, upon incorporation of these changes and another round of testing, the method will yield even more consistent and statistically relevant results.

First, none of the subjects at any point reached a steady-state condition. The subjects were all still increasing in metabolic rate at the completion of each task. It is highly desirable for data consistency purposes to have subjects reach or be close to steady state; as such, it is recommended that on further tests, the tasks are repeated for 4-5 minutes instead of the 2 minutes (or slightly higher) completed here.

Second, instead of having subjects repeat a task for a given amount of time, they should be asked to complete a set number of repetitions. This standardizes the amount of functional work they perform, which seems to be of value based on the results when normalized to the number of repetitions. As such, a number of repetitions for each task should be chosen such that the subject will require at least 4 minutes to complete the task. This minimum amount of time will ensure they get to or near steady state, as described above.

Third, after incorporating the first two changes, the variability within one test configuration needs to be characterized to determine if the differences between suits are relevant.

Fourth, the collection of unsuited data, while interesting for academic purposes, is not relevant when comparing one suit to another. In fact, no unsuited data are included in this report. Therefore, it could be possible to eliminate all unsuited runs from the test, which would allow for further reduced test time, subject fatigue, and overall test overhead. That said, familiarization of the tasks in the unsuited condition may be an important component and could be maintained with minimal impact. At the very least, it is recommended to choose only one method of unsuited task completion; i.e., the hybrid natural pace/fixed repetition method discussed above.

As for tasks, some changes are also warranted. Specifically:

- Walking: No changes warranted. Maintain 14' spacing to match rock panels used on reduced gravity plane, if desired.
- Sitting: Although this is a somewhat nonrepresentative EVA task, it serves as a good analog to a squat, which could be used for some surface exploration tasks. Therefore, no changes are warranted.
- Stair ascend/descend: No changes warranted.
- Prone and recover: Some subjects have difficulty completing this task; furthermore, it is not a task that a crewmember would nominally be required to do or that a suit would need to be optimized for. As such, it is recommended to change this task to a simple kneel and recover task, perhaps alternating knees to the floor or possibly in conjunction with picking up a small object off the floor. This will serve as a much more representative exploration task.
- Shoveling: It was determined that this task pushed grip strength and glove mobility/fit more so than overall suit mobility; therefore, it is recommended that this task be deleted until further developments warrant revisiting.
- Hammering: Although this task does work the arm bearing of a suit, it was determined that this task also became a test of grip strength for the subject rather than overall suit mobility. Therefore, it also is recommended for deletion until further developments warrant revisiting.
- Object relation: This task was easily completed by some subjects and barely completed by others; however, as it was considered primarily a fit issue, no changes are warranted at this time. Instead, in follow-up

testing, more importance should be placed on selecting subjects that fit the suit well; in addition, more emphasis and time should be placed on pre-test suit sizing modifications to optimize each subject's suit fit.

- Side step: No changes warranted.

Lastly, some general recommendations moving forward:

- Ensure subjects fit the suit very well. Subject B, for example, did not fit the Mark III extremely well and, as such, had issues in the prone and recover and object relocation tasks that made true comparison between suits difficult.
- The subject feedback in the form of rate of perceived exertion was not useful, as the difference between the two suits was always within one on a 1-10 scale. Furthermore, the queries between the two suits on a given task are hours apart, which makes subjective comparison more difficult. As a result, this portion of the test could be removed. However, based on increased test point durations as discussed above, it will be maintained until a more definite test protocol is determined.
- The Zephyr BioHarness data were not useful for data comparison purposes, especially as compared to the metabolic data. Although the Zephyr BioHarness data are of minimal impact to the test, they also could be removed with negligible loss.
- When possible, attempt to integrate an O₂ analyzer in addition to the CO₂ analyzer. This would eliminate the need to assume a constant respiratory exchange ratio and result in more accurate results.
- Use this pilot data to determine the necessary increase the sample size. At minimum, the number of subjects should increase to at least five, as we know there are at least that many people that have experience in these two suit designs and the necessary fitness level to complete the task.
- Measure subject fitness with a VO₂max test to determine both the relative and absolute intensity of these suited tasks.
- Look at the possibility of evaluating a suit in a different mass/mobility range. These two suits were comparable in both mass and mobility, and it would insightful to test a suit with lower mass (and, assumingly, corresponding lower mobility).
- Instead of the shoveling and hammering task, which are very glove-intensive, employ the use of an alternative upper-body task instead. One example might be a small board or wall where the subject moves objects from one place to another. This would tax different components of the suit than the object relocation task, as it would stress shoulder, arm, and – to a lesser extent – waist mobility by moving objects in a work envelope more commensurate with a microgravity EVA (i.e., waist to helmet height).

It should be noted that this test is assessing mobility in a 1-g environment when, in reality, suits are typically designed to perform best in microgravity or reduced-gravity environments. Although that is a shortcoming of this approach at this time, the goal is to develop a standard with low overhead that could be done anywhere. Therefore, it may be possible to later perform this mobility characterization technique in a more analogous gravity environment using the Active Response Gravity Offload System or something similar.

VIII. Conclusion

In conclusion, the approach of characterizing suited mobility through energy cost is very feasible. This pilot test demonstrated consistent, statistically relevant differences between suits on a per-task basis for several tasks. Note that this approach is currently evaluating only the differences between full suit configurations; the ability to extend this to the component level has yet to be evaluated or shown. Lessons learned should be incorporated into future rounds of testing to further improve the fidelity of the test method, with the ultimate goal of eventually using it as a tool at the suit engineer's disposal to evaluate spacesuit mobility with low overhead, low cost and high functional relevance.

References

¹Aitchison, L., "A Comparison of Methods for Assessing Space Suit Joint Ranges of Motion," AIAA-2012-3534, AIAA International Conference on Environmental Systems, 2012.

²Reinhardt, A., "Results and Applications of a Space Suit Range-of-Motion Study," SAE Technical Paper 891592, 1989, doi:10.4271/891592.

³Valish, D., Eversley, K. "A Comparison of Methods for Assessing Space Suit Joint Ranges of Motion," AIAA-2012-3532, AIAA International Conference on Environmental Systems, 2012.

⁴Matty, J. "Results and Analysis from Space Suit Joint Torque Testing," AIAA-2010-6211, AIAA International Conference on Environmental Systems, 2010.

⁵England, S., Benson, E., and Rajulu, S., "Functional Mobility Testing: Quantification of Functionally Utilized Mobility among Unsuites and Suited Subjects," NASA/TP-2010-216122, 2010.

⁶Aitchison, L., "Philosophies Applied in the Selection of Spacesuit Joint Range-of-Motion Requirements," *SAE Int. J. Aerosp.* 4(1):494-502, 2011.

⁷Norcross, J.R., Lee L.R., Clowers K.G., Morency R.M., Desantis L., De Witt J.K., et al. Feasibility of Performing a Suited 10-km Ambulation – Final Report of the EVA Walkback Test (EWT). NASA/TP-2009-214796, 2009.

⁸Norcross J.R., Clowers K.G., Clark T., Harvill L., Morency R.M., Stroud L.C., et al. Metabolic Costs and Biomechanics of Level Ambulation in a Planetary Suit. NASA/TP-2010-216115, 2010.

⁹Norcross J.R., Clowers K.G., Clark T., Harvill L., Morency R.M., Stroud L.C., et al. Metabolic Costs and Biomechanics of Inclined Ambulation and Exploration Tasks in a Planetary Suit. NASA/TP-2010-216215, 2010

¹⁰Chappell, S.P., Norcross, J.R., Abercromby, A.F., Gernhardt, M.L., "Life Sciences Implications of Lunar Surface Operations," NASA/TM-2010-216138, 2010.

¹¹McFarland, S. "Energy-Based Spacesuit Mobility Feasibility Assessment," NASA CTSD-ADV-1105, 2013.